

National Data Warehouse Project:

Data Integrity Verification Process



August 10, 2004

Office of Information Technology
Indian Health Service
Albuquerque, NM



NDW Data Integrity

The NDW is committed to assuring the integrity of your data throughout the NDW system.



What is “Data Integrity”

“Data Integrity means that data have not been altered or destroyed in an unauthorized manner... [It] is both a security and a quality principle that prevents information from being modified or otherwise corrupted, either maliciously or accidentally.”

National Committee on Vital and Health Statistics. Report to the Secretary of the U.S. Department of Health and Human Services on Uniform Data Standards for Patients Medical Record Information. July 2000



What is “Data Integrity”

... so, in simple terms, Data Integrity means that we will not make your good data bad, lose records, lose a diagnosis code, and so forth.

We should be able to show you that what we have at the NDW is exactly what you sent us.*

* Except for minor exceptions, e.g., non-Y2K compliant dates



What It is not...

Data integrity does not mean we can take poorer quality data and change it to make it “good quality.”

For example, data that is erroneous, missing, improperly coded, incomplete, etc. in the source systems will still be erroneous, missing, improperly coded, incomplete, etc. once it is loaded in the NDW.



"Data Quality"

The accuracy, completeness, appropriate use of codes, etc. of data in the source systems, its "data quality," has to be the responsibility of the local facilities who own and enter the data into their IT systems.

However we are partners. The NDW will provide local facilities information to assist them to improve the quality of the data in their systems.



NDW Data Quality Reports

- Field content checks against standard code tables
- Historical “norms” and deviation from norms
 - Records received versus expected
- Important missing data (e.g., SSNs)
- Timeliness of data



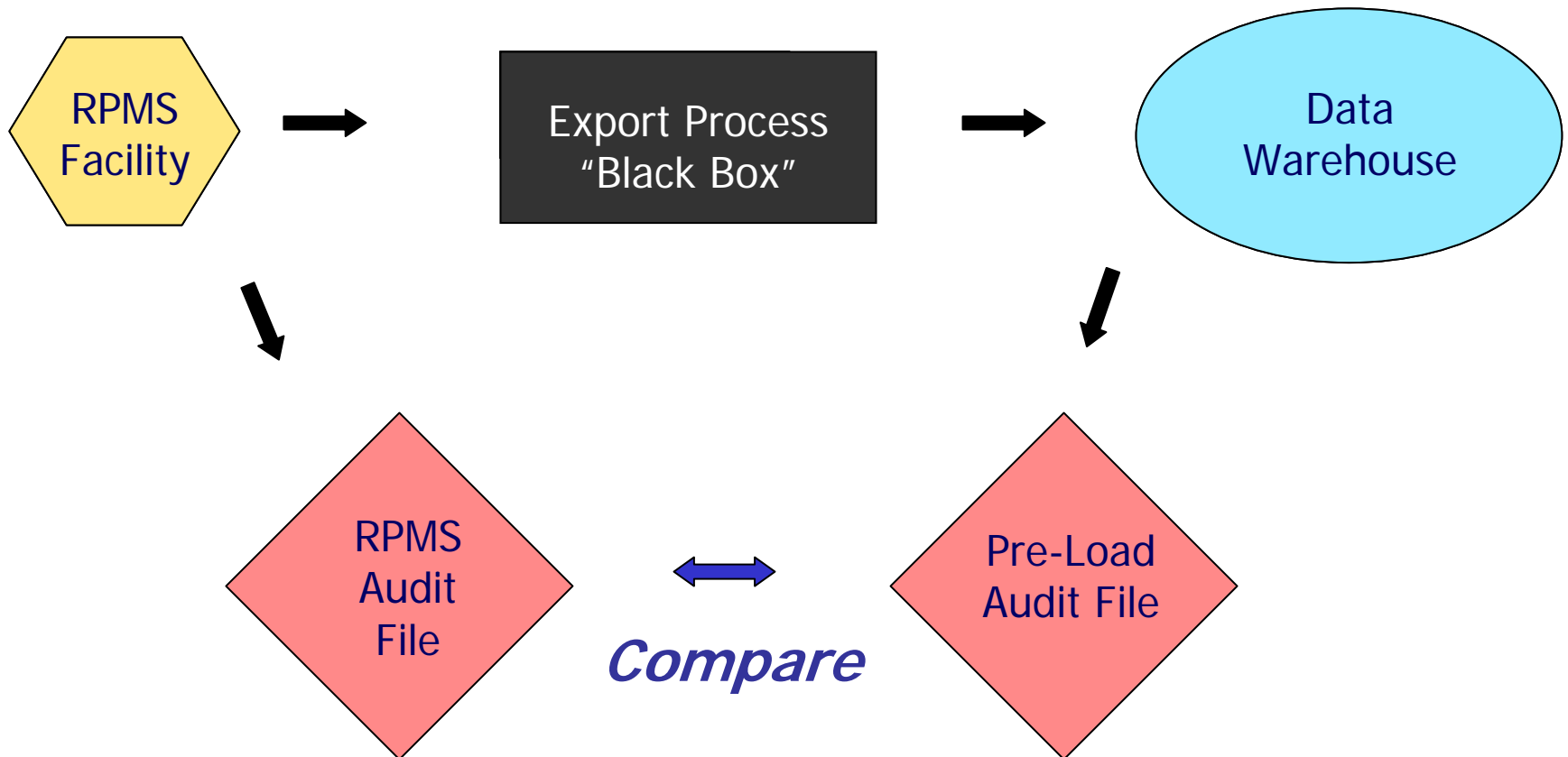
How Does the NDW Verify Data Integrity?

The NDW rigorously evaluates and monitors data integrity throughout the export of data, its routing to the NDW, and the load processes.

It does this in two steps:

- one that monitors the entire export and routing process; and*
- a second that monitors the load processes.*

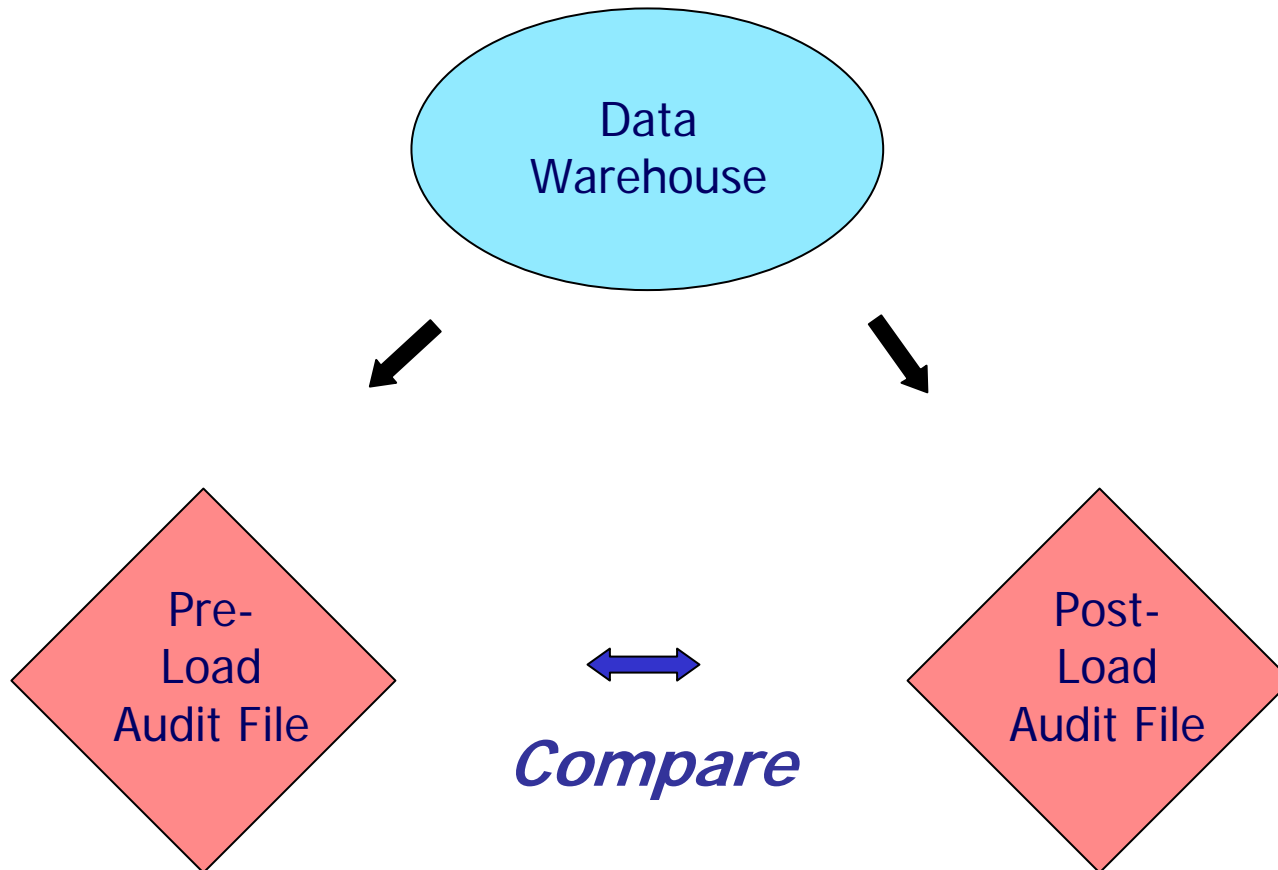
Data Integrity Verification Process – Step #1





Data Integrity Verification Process – Step #1

- At present the only export process we can monitor is IHS's RPMS export process
- But if non-RPMS sites send us appropriate audit files from their systems, we can do the same for them
- So we can even assist those with non-RPMS systems as they work to assure the integrity of their data through their proprietary exports!





What does the NDW Data Integrity Verification Process Monitor?

- Numbers of records
- Truncation of field entries
- Appropriate “mapping” of data elements from the source systems to the right place in the NDW system
- Serious errors that represent data corruption or processing errors



Summary Reports

- As each file export is received an acknowledgement is provided to the site
- After an Area's load is complete, a summary report will be produced
- These will allow our customers to notify us about any discrepancies they see.



Sample Acknowledgement

* DATA WAREHOUSE *
NPIRS Data Transmission Log Report
Data Received Date: Aug 06, 2004

For more information please send an email to: ITSCDataWarehouse@ihs.gov

Export ID: 130
Export file name: #####3040806134855.bdw.OUT
Static ASUFAC of exporting box: #####
Beginning Date: Oct 01, 2000
Ending Date: Mar 31, 2001
Run Location: SAMPLE SITE TRIBAL CLI
Transmission Status: No errors detected
Total Number of Encounter Records Received: 33245
Total Number of Registration Records Received: 0



Sample Area Summary Report

Under development...